Introduction to SAS[1]

The last of the great mainframe stats packages

---

[1]This slide show is an open-source document. See last slide for copyright information.

It almost seemed like there was one for every major university

- **DATATEXT**: Harvard
- **SPSS**: University of Chicago
- **BMDP**: University of California at Los Angeles
- **SAS**: University of North Carolina at Chapel Hill
- **OMNITAB**: Pennsylvania State University
- **S**: AT&T Bell labs

SAS *versus* R

- R is like a motorcycle.
- SAS is like a military Humvee.
- Except it doesn't break down.

SAS File Types
Not a complete list

- Raw data file
- Program file
- Log file
- Results file

- Data set also called Data Table
- Library

We work with these files

- **Raw Data File**: A file consisting of rows and columns of numbers; or maybe some of the columns have letters (character data) instead of numbers. The rows represent observations and the columns represent variables.
- **Program File**: A file consisting of commands that the SAS software tries to follow. You create this file with a text editor. The program file contains a reference to the raw data file (in the infile statement), so SAS knows where to find the data. Program files have names like `reading1.sas`.
- **Log File**: This file is produced by every SAS run, whether it is successful or unsuccessful. It contains a listing of the program file, as well any error messages or warnings. It will have a name like `reading1.log` or `reading1-log.html`.
- **Output File**: The list file contains the output of the statistical procedures requested by the program file. Output files have names like `reading1-Results.pdf`, `reading1-Results.rtf`, or `reading1-Results.html`.

SAS University Edition

- This is new.
- It seems to be the full version.
- It's free of charge to anybody with a university email address.

http://www.sas.com/en_us/software/university-edition.html

Features of SAS University Edition

- SAS lives in a virtual linux machine.
- You interact with it through a browser interface called SAS Studio.
- With SAS running in the virtual machine, you point your browser to a localhost address.
- The virtual linux machine is a web server.
- It looks like you are on the Internet, but everything is happening inside your computer.
- This way, it is really platform independent.
- You get your data into SAS via a shared folder, shared between the virtual linux machine and your physical machine.

More comments

- Its a big download – around 1.8 GB.
- Actually its two downloads.
- First, download the virtualization software, free from Oracle or VMWare. Do this first, *before* downloading SAS.
- Oracle's VirtualBox is recommended.
- The SAS download site has good instructions. Follow them *carefully*, because details matter.
- Once you connect to localhost, see the FAQ.
- Do this *before* starting the SAS software.
- The FAQ is actually a well-organized manual.
- Set up a shared folder according to instructions, and only then start SAS up for the first time.

Possible problems

- Slow or flaky internet connection. If it does not work the first time, try again.
- The virtual machine requires 1GB of RAM.
- Trouble with older operating systems?
- Not available in the computer labs.

- You can use SAS on `utstat` instead if you wish.
- There's also a version of SAS University Edition that runs on Amazon Web Services.

Copyright Information

This slide show was prepared by Jerry Brunner, Department of Statistics, University of Toronto. It is licensed under a Creative Commons Attribution - ShareAlike 3.0 Unported License. Use any part of it as you like and share the result freely. The LaTeX source code is available from the course website:

`http://www.utstat.toronto.edu/~brunner/oldclass/appliedf18`